INDUCED MATCHINGS AND INDUCED PATHS IN GRAPHS

JAROSLAV NEŠETŘIL AND PATRICE OSSONA DE MENDEZ

ABSTRACT. Denote by $\nabla_1(G)$ the maximum of $\frac{|E(H)|}{|V(H)|}$ over all (simple) minors of G obtained by contracting a star forest. We prove that there exists a positive function ϵ such that every graph G of order n has (at least) two clones (that is two vertices with the same neighbours) or an induced matching of size at least $\epsilon(\nabla_1(G))n$ and that this set may be found in linear time.

More generally, we prove that for every integer k there exists a (very slowly growing) positive function ϵ_k such that every graph of order n has an involutive automorphism or includes a set of size at least $k \lfloor \epsilon_k (\nabla_{\lfloor k/2 \rfloor}(G))n \rfloor$ inducing $\lfloor \epsilon_k (\nabla_{\lfloor k/2 \rfloor}(G))n \rfloor$ disjoint paths on k vertices.

1. INTRODUCTION

A matching of a graph G is a subset of pairwise non-adjacent edges. An *induced* matching of a graph G is a matching of G which is an induced subgraph of G, that is a matching with the property that no endpoint of an edge in the matching is adjacent to an endpoint of another edge in the matching.

The problem of finding a maximum induced matching (that is: an induced matching with maximum cardinality) has been introduced by Stockmeyer and Vazirani [39] as the "risk-free marriage problem" and it was studied extensively [13, 16, 17, 20, 38]. For a graph G we denote by $\beta^*(G)$ the size of a maximum induced matching.

It is known that the problem of deciding whether a given graph has an induced matching of size at least k (for given k) is NP-complete [39], even for bipartite graphs of maximum degree 4. However, this problem has been shown to be solvable in polynomial time for several graph classes [3, 4, 5, 6, 7, 19, 20, 24, 25, 26] and even in linear time for trees [18, 20, 42].

In this paper we consider the approximation version of this problem. Given a NP-complete optimization problem P, like the computation of the size of a maximum induced matching, it is usual to look for an approximation algorithm A_P such that the ratio of the cost of a feasible solution computed by A_P and the cost of an optimal solution is bounded by some constant R_{A_P} called the *performance ratio* of A_P . If P admits an approximation algorithm with performance ratio c, then we say that P is approximable within c. The class APX is the class of optimization problems that are approximable within c, for some constant c [2]. The approximation problem associated to the maximum induced matching problem consists in looking for an induced matching the size of which is at least within a factor c from the maximum. We say that P admits a polynomial time approximation scheme (PTAS) if, given any $\epsilon > 0$ there exists a polynomial-time approximation algorithm $A_{P,\epsilon}$ with performance ratio at most $1 + \epsilon$ [27]. An APX-complete optimization problem is an optimization problem which belongs to APX and to which any APX problem has an L-reduction in polynomial time (see [36] and [1] for a formal definition of an L-reduction). An important property of APX-completeness is that an APX-complete optimization problem Q does not admit a PTAS unless P = NP [1]. In particular, there is some constant c such that the problem of approximating Q within c is NP-hard.

Particularly, it is proved in [12] that in the class of *d*-regular graphs $(d \ge 3)$ the computation of $\beta^*(G)$ is approximable with asymptotic performance ratio (d-1) (hence belongs to APX) but is APX-complete. Here we shall extend the approximability result by proving that the problem of computing $\beta^*(G)$ is in APX when restricted to graphs with bounded $\nabla_1(G)$ which is defined (more generally) as follows:

Recall that the greatest reduced average degree (grad) with rank r of a graph G, denoted $\nabla_r(G)$, is defined by $\nabla_r(G) = \max \frac{|E(H)|}{|V(H)|}$, where the maximum is taken over all the minors H of G obtained by contracting a set of vertex-disjoint subgraphs with radius at most r and then deleting any number of edges and vertices [28, 29, 30, 32, 34]. Looking at the usual classes of sparse graphs, we see that minor closed classes satisfy $\nabla_r(G) < C$ for some constant C depending on the class, graphs with no subdivision of a fixed complete graphs (this includes classes of graphs with bounded degree) satisfy $\nabla_r(G) < f(r)$ for some function f depending on the class [32]. Thus these classes are more restrictive that the mere bounding of $\nabla_1(G)$, which in turn is more restrictive that a simple degeneracy condition (a graph G is k-degenerate iff $\nabla_0(G) \leq k/2$).

Our paper presents a further evidence that $\nabla_1(G)$ is an interesting parameter. In [31] we showed that graphs with bounded ∇_1 have linear Ramsey number and in [32] we showed a similar result for acyclic chromatic number.

In Section 4 we generalize the induced matchings to induced paths of length 2 and in Section 5 to induced paths of length k.

2. Definitions

The distance d(x, y) between two vertices x and y of a graph is the minimum length of a path linking x and y, or ∞ if x and y do not belong to the same connected component. The radius $\rho(G)$ of a connected graph G is the minimum maximum distance of the vertices from a fixed vertex, that is: $\rho(G) =$ $\min_{r \in V(G)} \max_{x \in V(G)} d(r, x)$. A vertex r is a center of G if the maximal distance of vertices of G to r is equal to $\rho(G)$. The radius $\rho(G)$ of a non-connected graph Gis the maximum of the radii of its components. We say that a graph G is F-free if G does not contain an induced subgraph isomorphic to F.

A (simple) graph H is a *minor* of a graph G if it may be obtained from G by contracting edges, deleting edges and deleting vertices. This is denoted by H < G. As edge deletions and contractions commute, we may consider contractions first and deletions next. As we only consider simple loopless graphs, each deletion is followed (if necessary) by the simplification of the graph. In other words, a minor H of a graph G is obtained by first contracting some connected subset F of edges, simplifying and then taking a subgraph. Symbolically we can write $H \subseteq G/F$ where G/F denotes the result of contracting of the set F of edges. Notice that the subset F is in general not uniquely determined by G and H. We denote by G_F the subgraph of G induced by the subset F of edges of G and by G[X] the subgraph induced by the subset X of vertices of G. The *depth* of a minor of a graph G is the minimum radius of the part we have to contract in G to get H. More formally we can write:

$$depth(H,G) = \min\{\rho(G_F) : H \subseteq G/F\}$$

Using this notation the definition of $\nabla_r(G)$ takes the following form:

Definition 2.1. ([28]) The greatest reduced average density (grad) of G with rank r is

$$\nabla_r(G) = \max_{\substack{H < G \\ \text{depth}(H,G) \le r}} \frac{|E(H)|}{|V(H)|}$$

If $\nabla_0(G)$ is closely related to the *coloring number* of G, $\nabla_1(G)$ is related to the *arrangeability* (used in the study of Ramsey numbers in [8]) and the *admissibility* (used in the study of game chromatic numbers in [22]), as shown in [31]. Also, $\nabla_{(k-1)/2}(G)$ is closely related to the *k*-coloring number introduced by Kierstead and Yang in [23], as shown by Zhu in [41]. We now state Zhu's result and the required preliminary definitions:

Let $<_L$ be a linear ordering of the vertices of a graph G, let k be an integer, and let $x <_L y$ be two vertices of G. The vertex x is weakly k-accessible from y if there exists an x-y path of length at most k in G, whose internal vertices are greater than x with respect to $<_L$. Denoting $Q_k(G, <_L, y)$ the set of the vertices which are weakly k-accessible from y, the weak k-coloring number wcol_k(G) is defined by:

$$\operatorname{wcol}_k(G) = 1 + \min_{<_L} \max_{v \in V(G)} |Q_k(G, <_L, y)|$$

where the minimum is taken over all the linear orderings of the vertex set V(G) of G. The following follows from [41]:

Lemma 2.1. Define $f_1(x) = 2x$ and for $i \ge 1$, $f_{i+1}(x) = f_1(x)f_i(x)^{2i^2}$. Then, for every graph G and every positive integer k, we have $\operatorname{wcol}_k(G) \le f_k(\nabla_{\lfloor k/2 \rfloor}(G))^k$.

We shall also make use of some results of [33]. We need some definitions. A rooted forest is a disjoint union of rooted trees. The height of a vertex x in a rooted forest F is the number of vertices of a path from the root (of the tree to which x belongs to) to x and is denoted by height(x, F). The height of F is the maximum height of the vertices of F. Let x, y be vertices of F. The vertex x is an ancestor of y in F if x belongs to the path linking y and the root of the tree of F to which y belongs to. The closure clos(F) of a rooted forest F is the graph with vertex set V(F) and edge set $\{\{x, y\} : x \text{ is an ancestor of } y \text{ in } F, x \neq y\}$. A rooted forest F defines a partial order on its set of vertices: $x \leq_F y$ if x is an ancestor of y in F. The comparability graph of this partial order is obviously clos(F). The tree-depth td(G) of a graph G is the minimum height of a rooted forest F such that $G \subseteq clos(F)$.

A centered coloring of a graph G is a vertex coloring such that, for any (induced) connected subgraph H, some color c(H) appears exactly once in H. This notion is similar to the ones of vertex ranking and ordered coloring which have been investigated in [10],[37]. As proved in [33], tree-depth and centered-coloring are closely related:

Lemma 2.2. Let G be a graph. Then, td(G) is the minimum number of colors in a centered coloring of G.

3. FINDING AN INDUCED MATCHING

A vertex v of a graph G is a *clone* if G has a vertex $u \neq v$ with the same neighbourhood as v. In that say we say that v is a *clone* of u. We denoted by \sim be the equivalence relation defined by $x \sim y$ if x and y have the same neighbors (i.e. are clones). Let G/\sim be the graph obtained by keeping exactly one vertex per equivalence class of \sim ; the vertex kept in a class is identified with the class it belongs to, so that for a vertex x of G belonging to a class represented by a vertex \hat{y} of G/\sim we write $x \in \hat{y}$. In a degenerated graph, the deletion of clones can only affect $\beta^*(G)$ by a constant factor: **Lemma 3.1.** Let G be a graph. Then G/\sim has no clones and the sizes of maximum induced matchings in G and G/\sim are related by

$$\beta^*(G) = \beta^*(G/\sim).$$

Proof. The fact that \sim is an equivalence relation is obvious. Moreover, it is straightforward that G/\sim has no clones: Assume \hat{x} and \hat{y} are vertices of G/\sim corresponding to the class of the vertices x and y of G. The neighbors of \hat{x} (resp. \hat{y}) are the classes of the neighbors of x (resp. y in G). If \hat{x} and \hat{y} have the same neighbors in G then x and y have the same neighbors in G hence belong to a same equivalence class of \sim . Thus the inequality $\beta^*(G) \geq \beta^*(G/\sim)$ is obvious as G/\sim is isomorphic to an induced subgraph of G.

For the second inequality, consider any maximum induced matching M of G. Let F be the set of edges of G/\sim defined by $\hat{x}\hat{y} \in F$ if there exists $x \in \hat{x}$ and $y \in \hat{y}$ such that $xy \in M$. Assume two edges of F are adjacent, namely $\hat{x}\hat{y}_1$ and $\hat{x}\hat{y}_2$. Then there exists $x_1, x_2 \in \hat{x}, y_1 \in \hat{y}_1, y_2 \in \hat{y}_2$ such that x_1y_1 and x_2y_2 belong to M. But y_1 is also a neighbor of x_2 (as it is a neighbor of x_1) hence M is not an induced matching, contradiction. It follows that F is an induced matching of G/\sim and $\beta^*(G) \leq \beta^*(G/\sim)$.

We shall now prove that a graph G of order n with no clones has an induced matching of size $\epsilon(\nabla_1(G))n$. To prove this, we will need two lemma which are inspired from the Lemma 2.1 and 2.2 of [35] (originally expressed in the particular case of graphs with no K_t minors, with aim of counting the number of graph of order n in a proper minor closed class) in a more precise form, related on the graph invariant $\nabla_0(G)$ (which is equivalent to the *degeneracy* or the *maximum average degree* of the graph up to a factor of 2). Recall that a clique in a graph G is any complete subgraph of G.

Lemma 3.2. Let G be a graph of order n. Let \mathcal{K}_t be the family of cliques of size t and let \mathcal{K} be the family of all cliques of G.

Then:

$$\begin{aligned} |\mathcal{K}_t| &\leq \binom{2\nabla_0(G)}{t-1}n\\ |\mathcal{K}| &\leq 2^{2\nabla_0(G)}n. \end{aligned}$$

The number of coverings of cliques by at most d (non-empty) sets is at most $2^{2\nabla_0(G)d}n$.

Proof. Consider an acyclic orientation of G with indegree at most $2\nabla_0(G)$. Then the vertices of any clique of size t are naturally ordered as x_1, x_2, \ldots, x_t (with all arcs oriented from x_i to x_j whenever i < j). We have at most n choices for x_t . The vertex x_t being given, we have at most $\binom{d^-(x_t)}{t-1}$ choices for $\{x_1, \ldots, x_{t-1}\}$. It follows the cardinality of \mathcal{K}_t is bounded by $\binom{\Delta^-(G)}{t-1}n = \binom{2\nabla_0(G)}{t-1}n$ and by summing we get $|\mathcal{K}| \leq 2^{2\nabla_0(G)}n$.

Covering a clique K of size t by at most d sets is equivalent to assigning to each vertex of K a non-empty subset of [d] (non-empty because the sets have to cover the clique) and then forgeting about the order of [d]. The result is simply $\sum_{t} {2\nabla_0(G) \choose t-1} n(2^d-1)^t/d!$.

Lemma 3.3. Let G be a graph of order n and let $d = 2\nabla_0(G)(2^{2\nabla_1(G)} + \nabla_1(G) + 1) + 1$. Then G has a subset S of vertices of size at least n/d such that every vertex in S has degree at most d and every vertex of S is either is a clone or it is adjacent to a vertex of degree at most d.

Proof. Let X be the set of all vertices of G that have degree at least d. Since the sum of the degrees of vertices of G is at most $2\nabla_0(G)n$, we get $|X| \leq \frac{2\nabla_0(G)n}{d}$. Let Y be the set of all vertices of V(G) - X that are adjacent to another vertex of V(G) - X. If $|Y| \ge n/d$, then Y satisfies the conclusion of Lemma 3.3, and so we may assume that |Y| < n/d. Let Z = V(G) - X - Y; then every neighbour of a vertex in Z belongs to X. In particular, no two vertices in Z are adjacent. Let Z' be a maximal subset of Z such that for every vertex $z \in Z'$ there exists a pair of distinct non-adjacent neighbours a(z), b(z) of z such that $\{a(z), b(z)\} \neq a(z), b(z)\}$ $\{a(z'), b(z')\}$ whenever $z, z' \in Z'$ are distinct. Let H be the graph obtained from $G[X \cup Z']$ by deleting, for each $z \in Z'$ all the edges incident to z but za(z) and zb(z)and then contracting za(z). As H is obtained from a subgraph of G by contracting a star forest, we have $\nabla_0(H) \leq \nabla_1(G)$. We notice that H may alternatively be defined as the graph obtained from G[X] by adding the edge a(z)b(z) for all $z \in Z'$. As there is a one-to-one mapping between the added edges and the vertices in Z'we deduce $|Z'| \leq |E(H)| \leq \nabla_0(H)|X| \leq \nabla_1(G)|X|$. The choice of Z' implies that the neighbourhood of every vertex of Z - Z' induces a clique of H. By Lemma 3.2 there are at most $2^{2\nabla_0(H)}|X| \leq 2^{2\nabla_1(G)}|X|$ of such cliques. In turn this implies that so all but possibly $2^{2\nabla_1(G)}|X|$ vertices of Z-Z' have degree at most d and are clones. But then we have

$$\begin{aligned} |Z| - |Z'| - 2^{2\nabla_1(G)} |X| &\ge n - |X \cup Y| - \nabla_1(G) |X| - 2^{2\nabla_1(G)} |X| \\ &\ge (1 - 2\nabla_0(G)(1 + \nabla_1(G) + 2^{2\nabla_1(G)})/d)n = n/d \end{aligned}$$

by the choice of d, as desired.

Theorem 3.4. Every connected graph G of order n has either a clone or an induced matching of size at least ϵn , where

$$\epsilon = \frac{1}{4\nabla_0(G)(2^{2\nabla_1(G)} + \nabla_1(G) + 1)(2\nabla_0(G)(2^{2\nabla_1(G)} + \nabla_1(G) + 1) + 1)^2}$$

Proof. Put $d = 2\nabla_0(G)(2^{2\nabla_1(G)} + \nabla_1(G) + 1) + 1$ (as in Lemma 3.3). Assume G has no clones. According to Lemma 3.3 there exists a subset S of vertices of size at least n/d such that every vertex in S has degree at most d and is adjacent to a vertex of degree at most d. For $x \in S$, let $\phi(x)$ be such an adjacent vertex of degree at most d. For $x \in S$ define $A(x) = N(x) \cup \{\phi(x)\} \cup N(\phi(x))$ and define B(x) as the union of the neighbourhoods of all the vertices in A(x) having degree at most d. Then $|A(x)| \leq 2d$ and $|B(x)| \leq 2d(d-1)$. Let $p = \frac{n}{2d^2(d-1)}$. Define iteratively $x_1, x_2, \ldots, x_i, \ldots, x_p$ by $x_i \in S \setminus \bigcup_{j < i} B(x_j)$. For $1 \leq i < j \leq p$ none of $x_j, \phi(x_j)$ may be adjacent to x_i or $\phi(x_i)$ by construction. Hence the set $\{x_1, \ldots, x_p, \phi(x_1), \ldots, \phi(x_p)\}$ forms an induced matching of size $p = \epsilon n$.

This result implies the desired *c*-approximation algorithm for $\beta^*(G)$:

Theorem 3.5. Let G be a connected graph. Then

$$\frac{\nabla_0(G)|V(G/\sim)|}{f(\nabla_0(G),\nabla_1(G))} \le \beta^*(G) \le \nabla_0(G)|V(G/\sim)|$$

where

$$f(x,y) = 4x^{2}(2^{2y} + y + 1)(2x(2^{2y} + y + 1) + 1)^{2}$$

Proof. The graph G/\sim has no induced matching of size greater than $|V(G/\sim)|/2$ but, by Theorem 3.4, it has an induced matching of size at least $\epsilon(\nabla_0(G), \nabla_1(G))n$. Choose $f(x, y) = x/\epsilon(x, y)$. Now the result follows from Lemma 3.3.



FIGURE 1. The subdivision graph $S(K_n)$ of K_n . Every edge of the maximum induced matching is incident to a principal vertex hence $\beta^*(S(K_n)) = n$.

Corollary 3.6. For every constant C, there exists a constant c > 0 such that an approximation of β^* in the class of graphs G with $\nabla_1(G) < C$ within a factor c may be computed in polynomial time, that is: MIM belongs to APX for bounded ∇_1 graphs.

On the one side, this results extends the proof by Zito [42] that *MIM* belongs to APX for *d*-regular graphs. On the other side, it is proved in [12] that the problem *MIM* is actually APX-complete in *d*-regular graphs for each $d \ge 3$. Thus there is some constant *c* such that the problem of *c*-approximating *MIM* for *d*-regular graphs is NP-hard unless P=NP.

Remark 3.7. Let S(G) denotes the subdivision graph of G (that is: the graph obtained from G by subdividing each edge exactly once). The graph $S(K_n)$ has no clones thus $S(K_n/\sim) = S(K_n)$. Moreover it is straightforward that $\beta^*(S(K_n)) = n$ (exactly one edge is incident to each principal vertex). Hence $\frac{|V(S(K_n)/\sim)|}{\beta^*(S(K_n))} = \frac{n+1}{2}$ although $\nabla_0(S(K_n)) < 2$. It follows that $\frac{|V(G/\sim)|}{\beta^*(G)}$ may not be bounded by a function of $\nabla_0(G)$ only. This shows that Theorem 3.5 is in this sense best possible.

4. FINDING INDUCED P_3 's

In this section we modify the results of Section 3 and obtain a similar result about ϵn copies of P_3 . We shall later see that this case shows the limitations of our method. But here we proceed similarly as in Section 3 and as in Lemma 3.3 we will now request for two small non-adjacent neighbors instead of a single small neighbor.

A non-identical automorphism α of a graph G is said to be *involutive* if α^2 is the identity. Clearly for every involutive automorphism α there exists a partition $V(G) = V_1 \cup V_2 \cup V_3$ such that $\alpha(x) = x$ for every $x \in V_1$ while $\alpha(x) \in V_3$ for every $x \in V_2$ and $\alpha(x) \in V_2$ for every $x \in V_3$. We say that α exchanges sets V_2 and V_3 . Note that by our assumption sets V_2 and V_3 are non-empty.

Lemma 4.1. Let G be a connected graph of order n with no involutive automorphism exchanging two cliques. Put

$$d = (1 + (2\nabla_0(G) + 1)\nabla_1(G) + 2^{2\nabla_1(G)(2\nabla_0(G) + 1) + 1})(2\nabla_0(G) + 1) + 1.$$



FIGURE 2. The subdivision graph $S(S_n)$ of the star S_n . An induced matching of size n is easily found but these graphs, although having no clones and bounded ∇_i for all $i \geq 0$, have no two disjoint induced P_3 .

Then G has a subset S of vertices of size at least n/d such that every vertex x in S has degree at most d and x has at least two non-adjacent neighbors which have degree at most d.

Proof. Let X be the set of all vertices of G that have degree at least d. Since the sum of the degrees of vertices of G is at most $2\nabla_0(G)n$, we get $|X| \leq \frac{2\nabla_0(G)n}{d}$. Let Y be the set of all vertices of V(G) - X which have at least two non-adjacent neighbors in V(G) - X. If $|Y| \ge n/d$, then Y satisfies the conclusion of the lemma, and so we may assume that |Y| < n/d. Let Z = V(G) - X - Y; then G[Z] is a disjoint union of cliques (of size at most $2\nabla_0(G) + 1$). Let Z' be a maximal subset of Z such that for every maximal clique $\omega \in G[Z']$ there exists a pair of distinct non-adjacent vertices $a(\omega), b(\omega) \notin Z$ each adjacent to some vertex of ω such that $\{a(\omega), b(\omega)\} \neq \{a(\omega'), b(\omega')\}$ whenever ω, ω' are distinct maximal cliques of G[Z']. Let H be the graph obtained from $G[X \cup Y \cup Z']$ by first contracting all the maximal cliques of G[Z'] (we denote by Z' the obtained vertex set and identify its elements with the maximal cliques of G[Z]; moreover we remove multiple edges), deleting, for each $\omega \in Z'$ all the incident edges but $\omega a(\omega)$ and $\omega b(\omega)$ and then contracting $\omega a(\omega)$. As H is obtained from a subgraph of G by contracting a star forest (consider the star formed by the neighborhood of a neighbor of $a(\omega)$ in the clique ω), we have $\nabla_0(H) \leq \nabla_1(G)$. We may notice that H may alternatively be defined as the graph obtained from $G[X \cup Y]$ by adding the edge $a(\omega)b(\omega)$ for all $\omega \in Z'$. As there is a one-to-one mapping between the added edges and the elements of Z' we deduce $|Z'| \le |E(H)| \le \nabla_0(H) |X \cup Y| \le \nabla_1(G) |X \cup Y|$ hence $|Z'| \le (2\nabla_0(G) + 1) |Z'| \le C_0(G) + C_0(G)$ $(2\nabla_0(G)+1)\nabla_1(G)|X\cup Y|$. The choice of Z' implies that the neighbourhood of every clique of G[Z - Z'] is a complete subgraph of H. By Lemma 3.2 there are at most $2^{2\nabla_0(H)(2\nabla_0(G)+1)+1}|X \cup Y| \leq 2^{2\nabla_1(G)(2\nabla_0(G)+1)+1}|X \cup Y|$ possible cliques and ways to connect their vertices to the vertices of a clique of size most $2\nabla_0(G) + 1$ in G[Z], and so if $|Z| - |Z'| > 2^{2\nabla_1(G)(2\nabla_0(G)+1)+1}|X \cup Y|$ two cliques of the same size are linked the same way to the same vertices of $X \cup Y$, contradicting the abscence

of such an automorphism of G. As

$$\begin{aligned} |Z| - |Z'| - 2^{2\nabla_1(G)(2\nabla_0(G)+1)+1} |X \cup Y| \\ \ge n - |X \cup Y| - (2\nabla_0(G)+1)\nabla_1(G)|X \cup Y| - 2^{2\nabla_1(G)(2\nabla_0(G)+1)+1} |X \cup Y| \\ \ge (1 - (1 + (2\nabla_0(G)+1)\nabla_1(G) + 2^{2\nabla_1(G)(2\nabla_0(G)+1)+1})(2\nabla_0(G)+1)/d)n \end{aligned}$$

and the contradiction follows from the choice of d.

Theorem 4.2. For every C > 0 there exists $\epsilon > 0$ such that every connected graph G of order n with no involutive automorphism exchanging two cliques and such that $\nabla_1(G) < C$ has a subset of $3\epsilon n$ vertices inducing ϵn disjoint paths of length 2.

Proof. Let d be defined as in Lemma 4.1. According to Lemma 4.1 there exists a subset S of vertices of size at least n/d such that every vertex in S has degree at most d and has at least two non-adjacent neighbors which have degree at most d. For $x \in S$, let $\phi_1(x), \phi_2(x)$ denote these two neighbors of x. For $x \in S$ define $A(x) = N(x) \cup \{\phi_1(x), \phi_2(x)\} \cup N(\phi_1(x)) \cup N(\phi_2(x))\}$ and define B(x) as the union of the neighbourhoods of all the vertices in A(x) having degree at most d. Then $|A(x)| \leq 3d - 1$ and $|B(x)| \leq 3d(d - 1)$. Let $p = \frac{n}{9d^2(d-1)}$. Define iteratively $x_1, x_2, \ldots, x_i, \ldots, x_p$ by $x_i \in S \setminus \bigcup_{j < i} B(x_j)$. For $1 \leq i < j \leq p$ none of $x_j, \phi_1(x_j), \phi_2(x_j)$ may be adjacent to x_i or $\phi_1(x_i), \phi_2(x_i)$ by construction. Hence $\{x_1, \ldots, x_p, \phi_1(x_1), \ldots, \phi_1(x_p), \phi_2(x_1), \ldots, \phi_2(x_p)\}$ induces $p = \epsilon n$ disjoints paths of length 2.

5. General case: Finding induced P_k 's

Finding induced P_k 's seems to be much more complicated than indicated by the cases k = 1, 2. Intuitively, the problem in generalizing the proof of Lemma 4.1 is that the connected components of G[Z], now having the weaker property to be induced P_k -free, cannot have their order bounded by a function of $\nabla_0(G)$. Of course, it would be easy to bound these orders by a function of d, but it would be of no help as it would eventually lead us to an implicite inequality which could be fullfilled by no choice of d. We have to proceede otherwise and invoke some more structure theory.

Lemma 5.1. Define $f_1(x) = 2x$ and for $i \ge 1$, $f_{i+1}(x) = f_1(x)f_i(x)^{2i^2}$. Let k be an integer and let G be a graph not containing P_k as an induced subgraph. Then $td(G) \le wcol_{k-2}(G) \le f_{k-2}(\nabla_{\lfloor k/2 \rfloor - 1}(G))^{k-2}$.

Proof. Consider a linear order L of the vertex set of G achieving $\operatorname{wcol}_{k-2}(G)$. Starting from the largest element of L, color successively the vertices of G such that any vertex x gets a color different from those vertices which are (k-2)-accessible from x. Obviously $\operatorname{wcol}_{k-2}(G)$ colors suffice.

Let H be any induced subgraph of G and let r be the smallest vertex of L. For each other vertex x of H, the shortest path in H from x to r has length at most k-2 and has r as its minimum. It follows that r is weakly (k-2)-accessible from x. Hence r has a color which is not present on another vertex of H. Thus the coloring is a centered coloring of G and, according to Lemma 2.2 $\operatorname{td}(G) \leq \operatorname{wcol}_{k-2}(G)$.

The bounding of wcol_{k-2}(G) by $f_{k-2}(\nabla_{\lfloor k/2 \rfloor - 1}(G))^{k-2}$ follows from Lemma 2.1.

We recall the following finitness theorem from [33]:

Theorem 5.2. There exists a function $F : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ with the following property: For any integer N, any graph G of order $n > F(N, \operatorname{td}(G))$ and any mapping $g : V(G) \to \{1, \ldots, N\}$, there exists a non trivial involuting g-preserving automorphism $\mu : G \to G$ with the fixed point property.



FIGURE 3. From a connected graph H with diameter k - 1, one can build a class with bounded expansion so that each graph in the class contains an induced P_k but not two disjoint ones.

Lemma 5.3. Let G be a graph having no non-trivial involuting automorphism. Let $A \subseteq V(G)$ be a subset of vertices and let H_1, \ldots, H_p be some of the connected components of G - A. Then

$$\sum_{1 \le i \le p} |V(H_i)| \le \mathcal{F}(|A|+1, |A| + \max_{1 \le i \le p} \operatorname{td}(H_i)).$$

Proof. Consider the coloring g of the subgraph G' of G induced by $A \cup \bigcup_{1 \le i \le p} V(H_i)$ such that all the vertices in A get a different color from the set $\{1, \ldots, |A|\}$ and all the other vertices get the color |A| + 1. It is obvious that any non-trivial involutive g-preserving automorphism of G defines a non-trivial involutive automorphism of G. According to Theorem 5.2, the graph G' has order at most $F(|A| + 1, \operatorname{td}(G'))$. By considering a vertex elimination order beginning by the deletion of the vertices in A we get $\operatorname{td}(G') \le |A| + \max_{1 \le i \le p} \operatorname{td}(H_i)$. Now the inequality $|V(G')| \ge \sum_{1 \le i \le p} |V(H_i)|$ finishes the proof. □

The necessity to consider the involutive automorphisms is indicated on Fig. 3.

Lemma 5.4. Let G be a connected graph of order n with no involutive automorphism φ exchanging two connected P_k -free subgraphs (this is schematically indicated on Fig. 4). Let k be an integer and put $d = (2\nabla_0(G) + 1)\left(1 + (k-1)\nabla_{\lfloor \frac{k}{2} \rfloor}(G) + 2^{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)}F\left(2\nabla_{\lfloor \frac{k}{2} \rfloor}(G) + 1 + (k-1)\left(2^{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G) + 1}\right), 2\nabla_{\lfloor \frac{k}{2} \rfloor}(G) + 1 + (k-1)\left(2^{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G) + 1}\right) + wcol_{k-2}(G)\right)$. Then G has a subset S of vertices of size at least n/d such that every vertex in S has degree at most d and belongs to an induced path of order k whose vertices all have degree at most d.



FIGURE 4. φ is an involutive automorphism exchanging two connected P_k -free subgraphs.

Proof. Let X be the set of all vertices of G that have degree at least d. Since the sum of the degrees of vertices of G is at most $2\nabla_0(G)n$, we get $|X| \leq \frac{2\nabla_0(G)n}{d}$. Let Y be the set of all vertices of V(G) - X which belong to an induced path of order k no vertices of which belong to X. If $|Y| \ge n/d$, then Y satisfies the conclusion of Lemma 5.4, and so we may assume that |Y| < n/d. Let Z = V(G) - X - Y; then G[Z] is a disjoint union of P_k -free graphs. Let \mathcal{P} be a maximal set internaly vertex sets disjoints induced paths of G of length at least 2, with endpoints in $X \cup Y$ and internal vertices in Z, such that $\{a(P), b(P)\} \neq \{a(P'), b(P')\}$ for any two distinct $P, P' \in \mathcal{P}$. Notice that every path in \mathcal{P} has length at most k. Let Z' be the union of the intersections of Z with the vertex sets of the paths in \mathcal{P} . Let H be the graph obtained from $G[X \cup Y \cup Z']$ by contracting all the paths in \mathcal{P} into single edges. As H is obtained from a subgraph of G by contracting disjoint balls of radius at most $\lceil \frac{k-1}{2} \rceil = \lfloor \frac{k}{2} \rfloor$, we have $\nabla_0(\hat{H}) \leq \nabla_{\lfloor \frac{k}{2} \rfloor}(G)$. We may notice that alternatively Hmay be defined as the graph obtained from $G[X \cup Y]$ by adding the edge a(P)b(P)for all $P \in \mathcal{P}$. As there is a one-to-one mapping between the added edges and the elements of \mathcal{P} we deduce $|\mathcal{P}| \leq |E(H)| \leq \nabla_0(H)|X \cup Y| \leq \nabla_{\lfloor \frac{k}{2} \rfloor}(G)|X \cup Y|$. Thus $|Z'| \le (k-1)\nabla_{\lfloor \frac{k}{2} \rfloor}(G)|X \cup Y|.$

The choice of \mathcal{P} implies that the neighbourhood of every connected component of G[Z - Z'] is a complete subgraph of H. By Lemma 3.2 there are at most $2^{2\nabla_0(H)}|X \cup Y| \leq 2^{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)}|X \cup Y|$ possible cliques, each of size at most $2\nabla_{\lfloor \frac{k}{2} \rfloor}(G) + 1$. Each of these cliques correspond in G to a subgraph of order at most $2\nabla_{\lfloor \frac{k}{2} \rfloor}(G) + 1 + (k-1)\binom{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1}{2}$. According to Lemma 5.3, we deduce that $|Z - Z'| \leq 2^{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)} \mathcal{F}\left(2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1+(k-1)\binom{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1}{2}, 2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1+(k-1)\binom{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1}{2} + wcol_{k-2}(G)\right)|X \cup Y|$. Summarizing, we get that the order of G is $n = |X \cup Y| + |Z'| + |Z - Z'| \leq \left(1 + (k-1)\nabla_{\lfloor \frac{k}{2} \rfloor}(G) + 2^{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)} \mathcal{F}\left(2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1+(k-1)\binom{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1}{2} + wcol_{k-2}(G)\right)\right)|X \cup Y| < \left(1 + (k-1)\nabla_{\lfloor \frac{k}{2} \rfloor}(G) + 2^{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)} \mathcal{F}\left(2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1+(k-1)\binom{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1}{2} + wcol_{k-2}(G)\right)\right)|X \cup Y| < \left(1 + (k-1)\nabla_{\lfloor \frac{k}{2} \rfloor}(G) + 2^{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)} \mathcal{F}\left(2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1+(k-1)\binom{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1}{2} + wcol_{k-2}(G)\right)\right)|X \cup Y| < (1 + (k-1)\nabla_{\lfloor \frac{k}{2} \rfloor}(G) + 2^{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)} \mathcal{F}\left(2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1+(k-1)\binom{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1}{2} + wcol_{k-2}(G)\right)|X \cup Y| < (1 + (k-1)(2^{2\nabla_{\lfloor \frac{k}{2} \rfloor}(G)+1) + wcol_{k-2}(G)))(2\nabla_0(G)+1)\frac{n}{d}$, which contradicts our choice of d. It follows that Y has order at least n/d.

Theorem 5.5. For every integer k > 2 and every C > 0 there exists $\epsilon > 0$ such that every connected graph G of order n with no involutive automorphism φ exchanging

two connected P_k -free subgraphs and such that $\nabla_{\lfloor k/2 \rfloor}(G) < C$ has a subset of ken vertices inducing en disjoint paths of order k.

6. Applications

6.1. Strong Star Chromatic Number. Dujmović and Wood introduced in [14] the following concept: a vertex coloring is a *strong star coloring* if between every pair of color classes, all edges (if any) are incident to a single vertex. That is, each bichromatic subgraph consists of a star and possibly some isolated vertices. The *strong star chromatic numver* of a graph G, denoted by $\chi_{sst}(G)$, is the minimum number of colors in a strong star coloring of G. The following result is proved in [14]:

Lemma 6.1. Every graph G with m edges and maximum degree $\Delta \ge 1$ has strong chromatic number $\chi_{\text{sst}}(G) \le 14\sqrt{\Delta m}$.

If the maximum degree is not bounded, the following upper bound is also proved in [14]:

Lemma 6.2. Every graph G with m edges has strong star chromatic number $\chi_{\rm sst}(G) \leq 15m^{2/3}$.

An asymptotically stronger result was known for graphs with bounded maximum degree in another context: a harmonious coloring of a simple graph G is a proper vertex coloring such that each pair of colors appears together on at most one edge. The harmonious chromatic number harm(G) is the least number of colors in such a coloring. Obviously, $\chi_{sst}(G) \leq harm(G)$. The following bounds are proved in [15]:

Theorem 6.3. Let Δ be a fixed integer, and $\epsilon > 0$. There is a natural number M such that if G is any graph with $m \geq M$ edges and maximum degree at most Δ , then the harmonious chromatic number harm(G) satisfies

$$\sqrt{2m} \le \operatorname{harm}(G) \le (1+\epsilon)\sqrt{2m}$$

The upper bound also applies to $\chi_{\rm sst}(G)$, but the derived lower bound for $\chi_{\rm sst}(G)$ is $\sqrt{\frac{2m}{\Delta}}$. Notice that in the most general case, we cannot expect any good lower bound for $\chi_{sst}(G)$ as the strong chromatic number of a star graph is 2. However, if we forbid clones we are able to prove a $\Theta(\sqrt{n})$ lower bound for graphs with bounded ∇_1 :

Theorem 6.4. Let G be a connected clone-free graph of order n and let

$$C(G) = 2\nabla_0(G)(2^{2\nabla_1(G)} + \nabla_1(G)).$$

Then the strong star chromatic number of G is bounded by

$$\chi_{\rm sst}(G) \ge \frac{1}{\sqrt{C(G)}(C(G)+1)}\sqrt{n}.$$

Proof. According to Theorem 3.4, the graph G has an induced matching of size $\frac{1}{2C(G)(C(G)+1)^2}n$. No two edges of the matching may have its endpoints colored by the same pair of colors. The result follows.

Remark 6.5. The condition that two color classes induce at most a star may be weakened to the condition that any two color classes induce a $2K_2$ -free bipartite graph. These graphs got alternative names and definitions like:

• *bipartite chain graphs* [40]: A graph is a bipartite chain graph if and only if it is bipartite and for each color class the neighbourhoods of the nodes in that color class can be ordered linearly with respect to inclusion (subset or equal);

- difference graphs [21]: A graph is a difference graph if every vertex v_i can be assigned a real number a_i and there exists a positive real number T such that (a) $|a_i| < T$ for all i and (b) $(v_i, v_j) \in E \iff |a_i a_j| \ge T$;
- non-separable bipartite graph [11]: A bipartite graph is non-separable if each pair of edges either share an end vertex or are connected by an edge.

Also, these graphs may be defined by the property that they are bipartite with at most one non-trivial connected component which is P_5 -free.

6.2. Extremal *H*-free graphs. It is known since [9] that the maximum number $ex^*(\Delta; H)$ of edges in a connected graph with maximum degree Δ and no induced subgraph *H* is finite if and only if *H* is a disjoint union of paths. It is also obvious that if the condition on the maximum degree is relaxed, no maximum exists. However we deduce the following from Theorem 5.5 (pP_k denotes the graph formed by *p* disjoint paths of length *k*):

Theorem 6.6. For every integer $k \geq 2$ and every real number D > 0, there exists a real number N(k, D) such that the maximum order of a P_k -free connected graph Gwith $\nabla_{\lfloor k/2 \rfloor}(G) < D$ and no involutive automorphism φ exchanging two connected pP_k -free subgraphs is at most pN(k, D).

Notice that we could not relax the absence of involutive automorphism exchanging two connected P_k -free subgraphs (consider the graphs of Fig 3). Also the the boundedness of a density parameter is needed as shown by the complement of any asymetric triangle-free graph. This graph has no induced P_5 .

Corollary 6.7. For every class C with bounded expansion there exists a function $N_{\mathcal{C}} : \mathbb{N} \to \mathbb{N}$, such that for every integers $k \geq 2$ and $p \geq 1$, the maximum number of edges in a connected pP_k -free graph $G \in C$ having no involutive automorphism is bounded by $pN_{\mathcal{C}}(k)$.

7. CONCLUSION

In this paper we have seen that the introduction of the invariant $\nabla_1(G)$ may be relevant to classical graph theoretical problems (see also [31] in the context of Ramsey numbers). We hope this will encourage the study of the classes of graphs with bounded ∇_r .

References

- G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi, Complexity and approximation, Springer-Verlag, Berlin, 1990.
- [2] _____, Complexity and approximation: Combinatorial optimization problems and their approximability properties., Springer-Verlag, Berlin, 1999.
- [3] A. Brandstädt, H.-O. Le, and J.-M. Vanherpe, Structure and stability number of chair-, co-Pand gem-free graphs revisited, Information Processing Letters 86 (2003), 161–167.
- [4] K. Cameron, Induced matchings, Discrete Applied Mathematics 24 (1989), 97–102.
- [5] _____, Induced matchings in intersection graphs, Discrete Mathematics 278 (2004), 1–9.
- [6] K. Cameron, R. Sritharan, and Y. Tang, Finding a maximum induced matching in weakly chordal graphs, Discrete Mathematics 266 (2003), 133–142.
- [7] J.-M. Chang, Induced matchings in asteroidal triple-free graphs, Discrete Applied Mathematics 132 (2004), 67–78.
- [8] G. T. Chen and R. H. Schelp, Graphs with linearly bounded ramsey numbers, Journal of Combinatorial Theory, Series B 57 (1993), no. 1, 138–149.
- [9] M. S. Chung and D. B. West, Large 2P₃-free graphs with bounded degree, Discrete Mathematics 150 (1996), no. 1-3, 69–79.
- [10] J.S. Deogun, T. Kloks, D. Kratsch, and H. Muller, On vertex ranking for permutation and other graphs, Proceedings of the 11th Annual Symposium on Theoretical Aspects of Computer Science (P. Enjalbert, E.W. Mayr, and K.W. Wagner, eds.), Lecture Notes in Computer Science, vol. 775, Springer, 1994, pp. 747–758.

- [11] G. Ding, Covering the edges with consecutive sets, Journal of Graph Theory 15 (1991), no. 5, 559–562.
- [12] W. Duckworth, D.F. Manlove, and M. Zito, On the approximability of the maximum induced matching problem, Journal of Discrete Algorithms 3 (2005), no. 1, 79–91.
- [13] W. Duckworth, N.C. Wormald, and M. Zito, Maximum induced matchings of random cubic graphs, Journal of Computational and Applied Mathematics 142 (2002), no. 1, 39–50.
- [14] V. Dujmović and D.R. Wood, Three-dimensional grid drawings with sub-quadratic volume, Proceedings of the 11th International Symposium on Graph Drawing (GD '03) (G. Liotta, ed.), Lecture Notes in Computer Science, vol. 2912, Springer-Verlag, 2004, pp. 190–201.
- [15] K. Edwards, The harmonious chromatic number of bounded degree graphs, Journal of London Mathematic Society 55 (1997), no. 2, 435–447.
- [16] P. Erdős, Problems and results in combinatorial analysis and graph theory, Discrete Mathematics 72 (1988), no. 1-3, 81–92.
- [17] R.J. Faudree, A. Gyárfás, R.H Schelp, and Zs. Tuza, *Induced matchings in bipartite graphs*, Discrete Mathematics **78** (1989), no. 1-2, 83–87.
- [18] G. Fricke and R. Laskar, Strong matchings in trees, Congressus Numerantium 89 (1992), 239–244.
- [19] M.C. Golumbic and M.C. Laskar, Irredundancy in circular arc graphs, Discrete Applied Mathematics 44 (1993), 79–89.
- [20] M.C. Golumbic and M. Lewenstein, New results on induced matchings, Discrete Applied Mathematics 101 (2000), no. 1-3, 157–165.
- [21] P. L. Hammer, U. N. Peled, and X. Sun, *Difference graphs*, Discrete Applied Mathematics 28 (1990), no. 1, 35–44.
- [22] H.A. Kierstead and W.T. Trotter, Planar graph coloring with an uncooperative partner, Journal of Graph Theory 18 (1994), no. 6, 569–584.
- [23] H.A. Kierstead and D. Yang, Orderings on graphs and game coloring number, Order 20 (2003), 255–264.
- [24] D. Kobler and U. Rotics, Finding maximum induced matchings in subclasses of claw-free and P₅-free graphs, and in graphs with matching and induced matching of equal maximum size, Algorithmica **37** (2003), 327–346.
- [25] V.V. Lozin, On maximum induced matchings in bipartite graphs, Information Processing Letters 81 (2002), 7–11.
- [26] V.V. Lozin and D. Rautenbach, Some results on graphs without long induced paths, Information Processing Letters 88 (2003), 167–171.
- [27] B.M. Moret, The theory of computation, Addison-Wesley, 1998.
- [28] J. Nešetřil and P. Ossona de Mendez, Grad and classes with bounded expansion I. decompositions, European Journal of Combinatorics (2005), (submitted).
- [29] _____, Grad and classes with bounded expansion I. decompositions, Tech. Report 2005-739, KAM-DIMATIA Series, 2005.
- [30] _____, The grad of a graph and classes with bounded expansion, 7th International Colloquium on Graph Theory (André Raspaud and Olivier Delmas, eds.), Electronic Notes in Discrete Mathematics, vol. 22, Elsevier, 2005, pp. 101–106.
- [31] _____, Fraternal augmentations, arrangeability and linearly bounded Ramsey numbers, Combinatorica (2006), submitted.
- [32] _____, Linear time low tree-width partitions and algorithmic consequences, STOC'06. Proceedings of the 38th Annual ACM Symposium on Theory of Computing, ACM Press, 2006, pp. 391–400.
- [33] _____, Tree depth, subgraph coloring and homomorphism bounds, European Journal of Combinatorics **27** (2006), no. 6, 1022–1041.
- [34] _____, Fraternal augmentations of graphs, coloration and minors, Proceedings of the Sixth Czech-Slovak International Symposium on Combinatorics, Graph Theory, Algorithms and Applications, Electronic Notes in Discrete Mathematics, vol. 28, 2007, pp. 223–230.
- [35] S. Norine, P. Seymour, R. Thomas, and P. Wollan, Proper minor-closed families are small, Journal of Combinatorial Theory, Series B 96 (2006), no. 5, 754–757.
- [36] C. H. Papadimitriou and M. Yannakakis, Optimization, approximation and complexity classes, Journal of Computer and System Sciences 43 (1991), 425–440.
- [37] P. Schaffer, Optimal node ranking of trees in linear time, Information Processing Letters 33 (1989/90), 91–96.
- [38] A. Steger and M.-l. Yu, On induced matchings, Discrete Mathematics 120 (1993), no. 1-3, 291–295.
- [39] L.J. Stockmeyer and V.V. Vazirani, NP-completeness of some generalizations of the maximum matching problem, Information Processing Letters 15 (1982), no. 1, 14–19.

- [40] M. Yannakakis, The complexity of the partial order dimension problem, SIAM Journal on Algebraic and Discrete Methods 3 (1982), 351–358.
- [41] X. Zhu, Colouring graphs with bounded generalized colouring number, Discrete Mathematics (2007), submitted.
- [42] M. Zito, Maximum induced matchings in regular graphs and trees, Proceedings of WG '99: the 25th International Workshop on Graph-Theoretic Concepts in Computer Science, Lecture Notes in Computer Science, vol. 1665, 1999, pp. 89–100.

DEPARTMENT OF APPLIED MATHEMATICS, AND, INSTITUTE OF THEORETICAL COMPUTER SCI-ENCE (ITI), CHARLES UNIVERSITY, MALOSTRANSKÉ NÁM.25, 11800 PRAHA 1, CZECH REPUBLIC *E-mail address*: nesetril@kam.ms.mff.cuni.cz

Centre d'Analyse et de Mathématiques Sociales, CNRS, UMR 8557, 54 B
d Raspail, 75006 Paris, France

 $E\text{-}mail \ address: \ \texttt{pom@ehess.fr}$